

PENGGOLONGAN SUARA BERDASARKAN USIA DENGAN MENGGUNAKAN METODE K – MEANS

Prabowo Hadi Putra Sutiknyo
Mahasiswa Politeknik Elektronika Negeri
Surabaya, Jurusan Teknik Telekomunikasi
Politeknik Elektronika Negeri Surabaya
Institut Teknologi Sepuluh Nopember
Kampus ITS, Surabaya 60111
e-mail : fuzink_aq@yahoo.co.id

ABSTRAK

Teknologi wicara adalah salah satu teknologi aplikasi yang telah ditemukan beberapa tahun lalu. Salah satunya adalah *speaker recognition* yang merupakan suatu proses yang sering disebut dengan verifikasi pengucap. Yang berarti mengenali suara dengan cara membandingkan dengan suara standar. Dengan mekanisme kerja pengambilan contoh-contoh suara, ekstraksi ciri dapat dilakukan dengan cara proses sampling, front-end, preemphasis, frame blocking, windowing, dan DFT (discrete Fourier Transform) dari ekstraksi ciri tersebut diproses lagi menggunakan metode *K-means* untuk mencari centroid diantara fitur-fitur dari hasil ekstraksi dan jarak terdekat antar centroid sehingga dapat dikelompokkan menjadi suara berdasarkan usia dewasa atau anak-anak yang kemudian disimpan ke dalam database. Pada saat ada sinyal wicara masuk, sistem akan melakukan proses pengolahan wicara. Kemudian hasil ekstraksi sinyal baru tersebut akan dibandingkan dengan hasil ekstraksi sinyal standar yang terdapat di database menggunakan metode *DFT* dan *K-means* sehingga akan dibandingkan dengan hasil pengklusteran, apakah suara tersebut masuk dalam range centroid 1 (dewasa) atau centroid 2 (anak-anak).

Hasil dari Software ini adalah berupa clustering suara dewasa dan anak-anak, yang mana nantinya sistem akan membedakan suara dewasa dan anak-anak dengan melihat nilai formant 2 nya.

Kata Kunci : speaker recognition, verifikasi pengucap, pengenalan wicara, pengelompokan jenis usia, Sampling, Front End Detection, Frame Blocking (20 ms), Pre – emphasis Filter, Windowing (Hammig), DFT, K – Means.

1. Pendahuluan

Data Clustering merupakan salah satu metode *Data Mining* yang bersifat tanpa arahan

(*unsupervised*). Ada dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu *hierarchical* (hirarki) data clustering dan *non-hierarchical* (non hirarki) data clustering. *K-Means* merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok. Metode ini mempartisi data ke dalam cluster/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain.

Adapun tujuan dari data clustering ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses clustering, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu cluster dan memaksimalkan variasi antar cluster. Data clustering menggunakan metode *K-Means* ini secara umum dilakukan dengan algoritma dasar sebagai berikut :

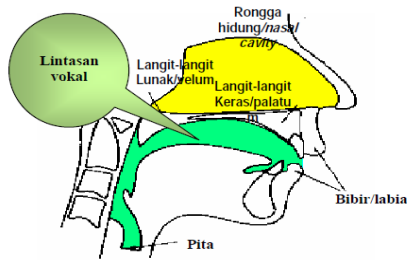
1. Tentukan jumlah cluster
2. Alokasikan data ke dalam cluster secara random
3. Hitung *centroid*/rata-rata dari data yang ada di masing-masing cluster
4. Alokasikan masing-masing data ke *centroid*/rata-rata terdekat
5. Kembali ke Step 3, apabila masih ada data yang berpindah cluster atau apabila

perubahan nilai *centroid*, ada yang di atas nilai *threshold* yang ditentukan atau apabila perubahan nilai pada *objective function* yang digunakan di atas nilai *threshold* yang ditentukan.

2. Teori Penunjang

2.1 Sinyal Suara Manusia

Manusia menggunakan suara sebagai sumber informasi untuk mengkomunikasikan ide, keinginan, dan perasaannya kepada orang lain. Organ tubuh yang berpengaruh dalam proses produksi wicara adalah paru-paru, tenggorokan (trachea), larinks, farinks, rongga hidung (nasal cavity), dan rongga mulut (oral cavity). Pembangkitan sinyal suara terletak pada bentuk lintasan vokalnya (vocal tract). Lintasan vokal tersebut terdiri atas: dibawah katub tenggorokan (laryngeal pharynx), antara langit-langit lunak katub tenggorokan (oral pharynx), di atas velum dan diujung depan rongga hidung (nasal pharynx), dan rongga hidung (nasal cavity), seperti ditunjukkan pada gambar di bawah ini:



Gambar 1

2.2 Proses Sampling

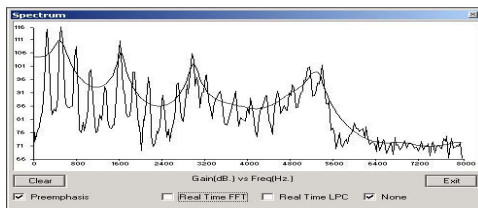
Sinyal suara merupakan sinyal yang tidak terbatas dalam domain waktu (infinite time interval). Suara manusia akan menghasilkan sinyal analog yang terus kontinu. Untuk keperluan pemrosesan dalam transformasi fourier maka sinyal wicara harus dibentuk dalam potongan potongan waktu yang terbatas (finite time interval). Karena itu sinyal yang ada dipotong-potong dalam slot-slot interval waktu tertentu.

2.3 Front End Detection

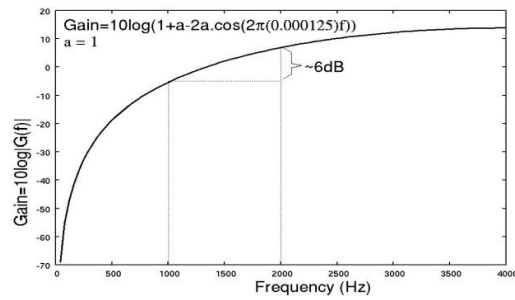
Front End pada speech recognition digunakan untuk mendeskripsikan sistem dimana mesin mendeteksi ke dalam sistem speech recognition, pengenalan kata akan didisplaykan secara benar setelah berbicara, dan pendeteksi memungkinkan untuk mengkoreksi beberapa bagian yang tidak dikenali (*unvoiced*, dan *noise*) pada file tersebut. Bagian-bagian ini tidak akan bisa lewat melalui pengeditan, dalam istilahnya.

2.4 Pre emphasis

untuk mengurangi efek ketelitian terbatas dalam proses pemfilteran pre-emphasis sistem adaptif tetap, atau berubah perlahan (Untuk kondisi transmisi rata-rata, latar belakang derau, atau bahkan meratakan spectrum sinyal). Pre-emphasis yang banyak digunakan adalah sistem berderajat pertama.



Gambar 2 Respon frekuensi preemphasis

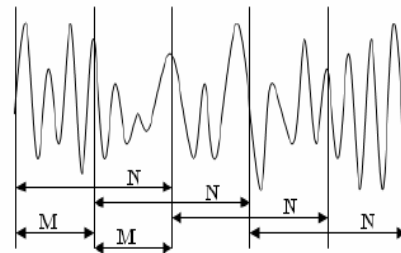


Gambar 3 Bentuk representasi dari sinyal suara setelah melalui proses formant dan preemphasis

2.5 Frame Blocking

Pada langkah ini, sinyal wicara yang telah dipreemphasis, $s(n)$, diblok dalam frame-frame dengan N sample dan digeser sebesar M sample dimana $N \approx 2 \times M$, seperti ditunjukkan pada gambar 4 Sehingga didapatkan nilai dari sinyal yang baru adalah sebagai berikut :

$$x_l(n) = s(M \cdot l + n), \quad n = 0, 1, 2, \dots, N-1 \quad l = 0, 1, 2, \dots, L-1 \quad (1)$$



Gambar 4 Bentuk sinyal yang diframe blocking

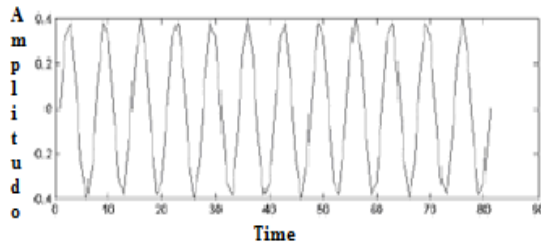
2.6 Windowing

Pada langkah ini dilakukan fungsi *weighting* pada setiap frame yang telah dibentuk pada langkah sebelumnya. Windowing diperlukan untuk mengurangi efek diskontinuitas dari potongan – potongan sinyal.

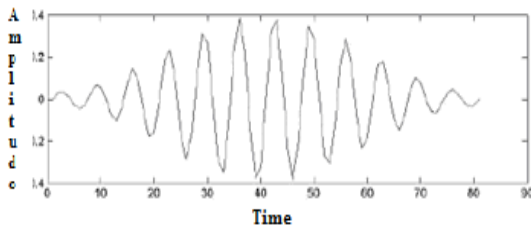
Dimana jenis windowing ada beberapa macam yaitu Hamming, Hanning, Bartlet, Rectanguler dan Blackman. Namun dalam penelitian ini menggunakan Window Hamming dikarenakan hasilnya lebih halus.

♦ Window Hamming

$$W_{ham}(n) = \begin{cases} 0.52 - 0.46 \cos[\pi n / (N-1)] & 0 \leq n \leq N-1 \\ 0 & \text{elsewhere} \end{cases}$$



Gambar 5 Sinyal sinus murni

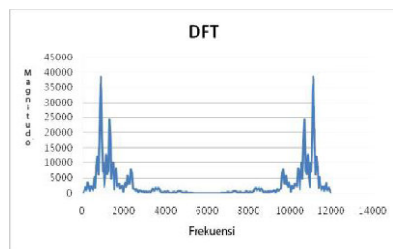


Gambar 6 sinyal sinus yang telah di window

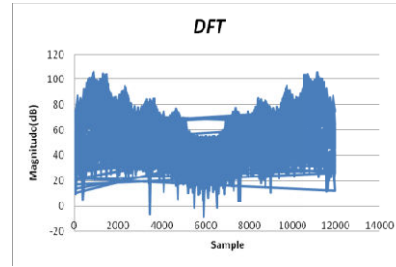
2.7 DFT (Discrete Fourier Transform)

Algoritma yang digunakan untuk mengubah sampel data dari domain waktu ke domain frekuensi adalah Discrete Fourier Transform (DFT). DFT menstabilkan hubungan antara sampel – sample signal domain waktu dan merepresentasikannya ke domain frekuensi. DFT biasa digunakan dalam analisis spectral, diaplikasikan dalam mekanik, akustik, pencitraan medis, analisa angka, instrumentasi dan telekomunikasi.

Yang mana DFT dalam bentuk domain frekuensi dapat dilihat pada gambar dibawah :



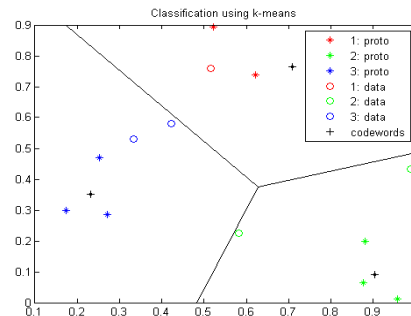
Gambar 7 bentuk DFT dalam domain frekuensi



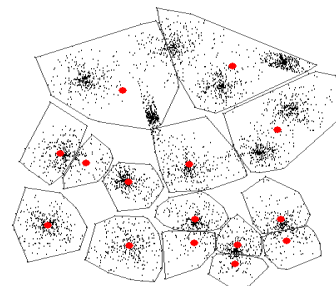
Gambar 8 bentuk DFT setelah dilogkan

2.8 Algoritma K-Means

Algoritma K-Means merupakan satu algoritma yang sering kali digunakan di dalam teknik pengelompokan karena membuat suatu perkiraan yang efisien dan tidak memerlukan banyak parameter. K-Means [MacQueen(1967)] menggunakan k kelompok yang telah ditetapkan (k kelompok pertama sebagai centroid). seperti pada gambar 9 dan 10, secara berterusan akan melalui proses perkiraan titik tengah (jarak minimum) sehingga sesuatu fungsi kriteria dicapai (kelompok adalah tetap).



Gambar 9 Contoh perolehan centroid menggunakan K-means pada matlab



Gambar 10 Contoh perolehan centroid menggunakan K-means

2.8 Tcl

Tcl merupakan singkatan dari *Tool Control Language*. Sedangkan *Tk* adalah *Graphical Toolkit extension* dari *Tcl*. *Tcl/tk* menyediakan bermacam-macam item standar antarmuka *GUI* untuk memfasilitasi user untuk membuat sebuah tampilan atau desain secara cepat dan juga bisa untuk pengembangan aplikasi tingkat tinggi lainnya.

Untuk bahasa pemrograman di *Tcl/tk* bentuknya sama seperti bahasa pemrograman di C/C++ terutama pada *Loop structures*, definisi fungsi dan logika penghitungannya. Catatan, didalam *Tcl* semua data di representasikan sebagai string.

2.10 Snack

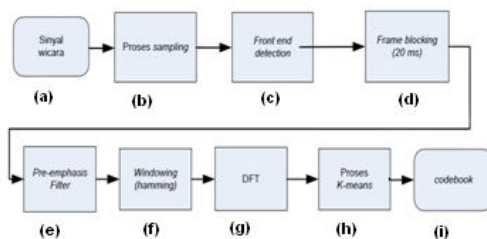
Snack merupakan software yang digunakan untuk pemrosesan suara yang dipakai sebagai *extensi* dalam suatu bahasa pemrograman. *Snack* biasanya digunakan bersama-sama dengan bahasa pemrograman script seperti *Tcl/Tk*, *Python*, *Ruby*. Dalam proyek akhir ini akan menggunakan software *snack* versi 2.2.2.n



Gambar 11 Contoh aplikasi Snack

3. Perancangan System

Pada Bab 3 ini akan dijelaskan tentang perancangan Sistem mulai dari pengambilan data berupa sample suara sampai pada pengelompokkan atau *clustering* dengan menggunakan metode *K – Means* Diagram blok untuk pengolahan sinyal suara pada fase pembelajaran adalah sebagai berikut:



Gambar 12 Blok Diagram Fase Pembelajaran

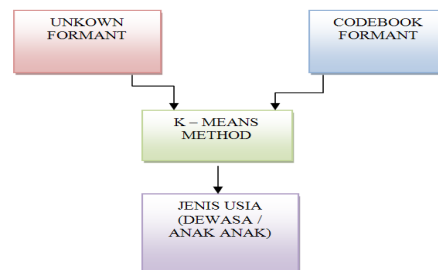
Yang mana penjelasan proses dari gambar 12 adalah sebagai berikut :

- Sinyal Wicara direkam dengan Audacity
- Setelah direkam dan disimpan dalam ekstensi .wav proses berikutnya adalah dilakukan proses Sampling
- Setelah Melewati Proses Sampling Masuk ke proses Front End Detection
- Kemudian Masuk lagi ke proses Frame Blocking (20 ms)
- Diteruskan ke Proses Pre – emphasis Filter
- Dilanjutkan lagi ke proses Windowing (hamming)
- Kemudian dilanjutkan ke proses DFT
- Setelah mendapatkan nilai Formantnya dari proses DFT di kluster dengan Proses K – Means
- Setelah melewati proses K – Means data di simpan dalam Code book sebagai data pembanding.

Semua program sistem pengolahan suara dibuat dalam Bahasa Pemrograman C dengan kompiler Microsoft visual C++ 6.0 dan dibuat dalam file berekstensi .dll agar nantinya dapat diintegrasikan dengan *Tcl/Tk* dan *Snack*.

Fase pengujian

Pada tahap pemrosesan suara didapatkan hasil berupa *formant* 1 dan *formant* 2. Yang mana Nilai *formant* diperoleh dari nilai *magnitude* tertinggi atau puncak dari suatu *frame* antara rentang frekuensi tertentu yang dapat mewakili suara pengguna. Dari data inilah yang digunakan sebagai *codebook*. Pada saat pengujian, suara tidak diketahui (*unknown voiced*) dibandingkan dengan 2 buah *centroid* yang mewakili suara dewasa dan anak-anak dengan menghitung jarak minimum (*Euclidean Distance*) sehingga oleh sistem dapat diputuskan suara tidak diketahui tersebut sebagai suara dewasa atau anak-anak. Proses pembandingan pada waktu proses pengujian dapat dilihat pada gambar 13 :



Gambar 13 Fase Pengujian suara

- Unknown Formant**

Unknown Formant adalah hasil formant dari suara yang belum diketahui. Pada proyek akhir ini hasil formant yang ingin

dikenali disimpan dalam file berbentuk “formant.txt”

- **Codebook Formant**

File “formant.txt” mewakili dari *magnitude* tertinggi yang disebut sebagai *formant* dari 50 pewicara yang disimpan dalam *codebook* atau *database*.

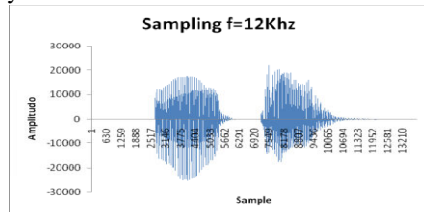
4. Hasil Kemajuan proyek akhir

Pembahasan pada paper ini dititik beratkan pada DFT (*discrete fourier Transform*) dan *K – means* dimana pada proses tersebut akan didapatkan nilai Formant yang nantinya akan dicluster dengan menggunakan metode *K – Means*. Namun untuk mendapatkan hasil tersebut terlebih dahulu melalui beberapa proses yakni sebagai berikut :

4.1 Proses Sampling

Pada Proses Awal dari pembuatan software pengenalan suara ini adalah program sampling. Yang mana persyaratan frekuensi sampling menurut teorema *Shannon* harus sama dengan atau melebihi 2 kali frekuensi sinyal yang di sample.

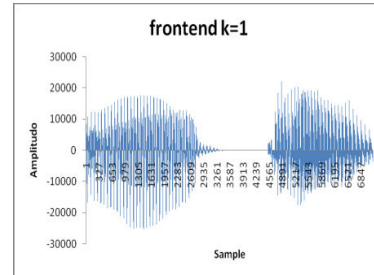
Pada system ini menggunakan $F_s = 12$ pada pada gambar 14 karena frekuensi tersebut telah sesuai dengan teori sampling *Nyquist* yaitu $f_{sampling} \geq 2 f_{sinyal}$ jelas terlihat bahwa sampling telah merepresentasikan sinyal asli sehingga dapat dipergunakan untuk proses selanjutnya.



Gambar 14 Proses Sampling dengan $f = 12\text{KHz}$

4.2 Front End Detection

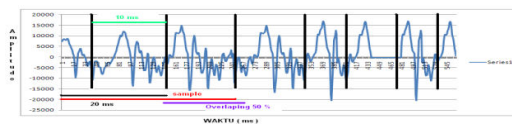
Dalam proyek akhir ini menggunakan harga $k=1$ agar dapat dilanjutkan ke proses selanjutnya untuk mendapatkan fitur yang diinginkan. Dikarenakan dengan menggunakan nilai konstanta 0.2 atau < 1 masih terdapat Noise diawal dan akhir sinyal Informasi. Sedangkan pada Nilai konstanta 2 sampai 4 atau > 1 sinyal informasi ikut terpotong yang menyebabkan sinyal informasi yang dibutuhkan semakin berkurang.



Gambar 15 Proses Front end $k = 1$

4.3 Frame Blocking

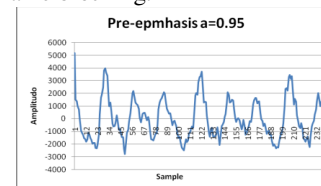
Proses selanjutnya yaitu proses frame blocking untuk mempermudah dalam penganalisaan sinyal pada proses pengambilan fitur. Pada proses ini dilakukan pemotongan sinyal dalam slot-slot tertentu agar memenuhi 2 syarat yaitu linear dan time invariant. Pada proyek akhir ini sinyal suara dipotong sepanjang 20 milidetik. Setiap potongan tersebut disebut frame. Jadi dalam satu frame terdapat 240 sampel dari 12000 sampel yang ada. Hasil nilai dari proses ini adalah sebagai berikut:



Gambar 16 Proses Frame blocking dengan overlapping 50%

4.4 Pre – Emphasis

Pada proses pre-emphasis sinyal dengan frekuensi tinggi di loloskan dan sinyal dengan frekuensi rendah difilter hal ini dilakukan dengan tujuan untuk meratakan bentuk sinyal untuk frekuensi tinggi maupun rendah suara seseorang akibat penurunan nilai daya kuat suara sebesar -6 dB/oktaf pada waktu dikeluarkan dari bibir setelah proses frame blocking.

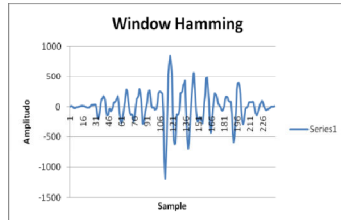


Gambar 17 proses Pre emphasis $a = 0,95$

Pre-emphasis $a = 0.95$ sinyal informasi hampir rata secara keseluruhan sehingga sinyal dapat dilakukan ke proses selanjutnya. Sehingga dapat disimpulkan bahwa sinyal dari proses Frame Blocking yang telah dilakukan proses pre-emphasis, sinyal pada frekuensi rendah telah di filter.

4.5 Windowing

Proses berikutnya adalah Proses Windowing, proses windowing ini untuk mengurangi efek diskontinuitas ketika sinyal ditransformasikan ke domain frekuensi.



Gambar 18 proses windowing hamming

Dari gambar 18 dapat dilihat bahwa pada Sinyal yang telah melewati proses *Window Hamming* pada awal dan akhir sinyal lebih halus daripada sinyal sebelumnya, sehingga tidak terjadi efek *discontinuitas* atau sinyal seolah – olah terputus.

4.6 DFT (Discrete Fourier Transform)

Pada proses DFT sinyal dirubah dari domain waktu ke domain frekuensi. Dan menghasilkan nilai magnitude. Dari nilai magnitude itu kemudian dihasilkan nilai formant 1 dan formant 2.

- Untuk mendapatkan fitur – fitur sinyal dari proses *DFT* maka perlu dilakukan proses smoothing dari nilai magnitude yang didapat, agar sinyal lebih terlihat halus kemudian diambil puncak – puncak sinyal, formant yang diambil sebagai perwakilan suara adalah formant1 dan formant2 seperti pada tabel 2 dan 3.
- Formant merupakan warna suara yang digambarkan dalam frequency domain. Suara manusia memiliki karakteristik Formant tertentu yang apabila dirubah akan menyebabkan perubahan warna suara.

Tabel 1 Nilai magnitude dari proses DFT dalam

Frekuensi (Hz)	Magnitude
0.000000	76.085
46.000000	741.499
93.000000	94.309
140.000000	104.903
187.000000	100.368
234.000000	887.053
281.000000	102.336
328.000000	100.979
375.000000	871.599
421.000000	105.39
468.000000	105.357
515.000000	842.787
562.000000	107.037
609.000000	109.938
656.000000	986.792
703.000000	946.704
750.000000	102.606
796.000000	935.544
843.000000	106.369
890.000000	111.214

937.000000	102.258
984.000000	987.571
1.031.000000	107.199
1.078.000000	102.105
1.125.000000	835.292
1.171.000000	893.218
.....

Tabel 2 hasil dari proses smoothing karakteristik suara (Dewasa)

Data Ke	Formant 1	Formant 2
Data ke 1	421.000000	3093.000000
Data ke 2	562.000000	3187.000000
Data ke 3	515.000000	3093.000000
Data ke 4	562.000000	3093.000000
Data ke 5	890.000000	3187.000000
Data ke 6	750.000000	3000.000000
Data ke 7	562.000000	3187.000000
Data ke 8	515.000000	3187.000000
Data ke 9	703.000000	3187.000000
Data ke 10	703.000000	2625.000000
Data ke 11	843.000000	3093.000000
Data ke 12	234.000000	2906.000000
Data ke 13	46.000000	2859.000000
Data ke 14	750.000000	2671.000000
Data ke 15	843.000000	2953.000000
Data ke 16	796.000000	2812.000000
Data ke 17	890.000000	2859.000000
Data ke 18	46.000000	3187.000000
Data ke 19	843.000000	2812.000000
Data ke 20	609.000000	2765.000000
Data ke 21	328.000000	3140.000000
Data ke 22	515.000000	2625.000000
Data ke 23	46.000000	2765.000000
Data ke 24	843.000000	2812.000000
Data ke 25	796.000000	3140.000000

Tabel 3 hasil dari proses smoothing karakteristik suara (anak – anak)

Data Ke	Formant 1	Formant 2
Data ke 1	984.000000	2015.000000
Data ke 2	609.000000	2062.000000
Data ke 3	281.000000	2437.000000
Data ke 4	468.000000	2250.000000
Data ke 5	937.000000	2062.000000
Data ke 6	843.000000	2015.000000
Data ke 7	656.000000	2156.000000
Data ke 8	750.000000	2062.000000
Data ke 9	843.000000	2015.000000
Data ke 10	281.000000	2062.000000
Data ke 11	515.000000	2015.000000
Data ke 12	609.000000	2109.000000
Data ke 13	984.000000	2203.000000
Data ke 14	515.000000	2296.000000
Data ke 15	703.000000	2109.000000
Data ke 16	468.000000	2015.000000
Data ke 17	984.000000	2156.000000
Data ke 18	750.000000	2062.000000
Data ke 19	375.000000	2343.000000

Data ke 20	937.000000	2109.000000
Data ke 21	375.000000	2062.000000
Data ke 22	375.000000	2296.000000
Data ke 23	984.000000	2015.000000
Data ke 24	609.000000	2296.000000
Data ke 25	937.000000	2203.000000

Untuk penjelasanya adalah sebagai berikut

- Dari hasil *DFT* diperoleh nilai formant tiap penutur atau pewicara. Jumlah penutur sebanyak 50 orang, yang terdiri dari 25 Dewasa dan 25 Anak anak, tetapi karena terdapat data yang dinyatakan layak yaitu data yang membentuk 2 kelompok sebanyak 36 data, sisanya yaitu sebanyak 14 buah data dari kelompok dewasa maupun anak anak adalah data yang mempunyai nilai yang sama. Pada proyek akhir ini kelas Dewasa disimbolkan dengan “● **Lingkaran merah**” sedangkan kelas Anak anak sebagai “● **Lingkaran biru**”.

4.7 Pengujian dan Analisa K – Means

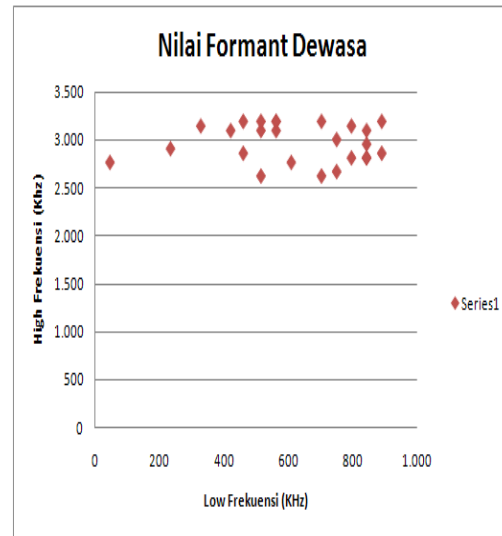
Seperti yang telah dijelaskan pada teori sebelumnya Pada Proyek Akhir ini untuk menghasilkan pengenalan berdasarkan jenis usia, maka pembagian kelompok dibagi menjadi 2 kelas yakni kelompok Dewasa dan Kelompok Anak – Anak. Dari nilai *formant* yang diperoleh secara keseluruhan, untuk diperoleh hasil berupa penggolongan Dewasa atau Anak langkah – langkahnya adalah sebagai berikut :

1. langkah pertama dari proses *k-means* adalah menentukan *centroid* setiap kelompok. Dengan kondisi kelompok pertama menggunakan 1 *formant* sebagai *centroid* sedangkan *centroid* yang lain adalah hasil rata-rata seluruh formant yang ada.
2. Kemudian langkah kedua adalah menghitung jarak *Euclidean* setiap *formant* terhadap 2 *centroid* yang berbeda, maka akan diperoleh nilai *Euclidean* yang berbeda untuk setiap *formant* terhadap 2 *centroid* tersebut. Dari perolehan data tersebut dapat dilihat nilai terkecil *Euclidean* setiap *formant* terhadap kedua *centroid* tersebut hingga dapat dikelompokkan *formant* tersebut anggota dari *centroid* 1 atau 2.

Pengelompokan tersebut dicari setiap *centroid* dengan jumlah *formant* yang berbeda secara berulang-ulang terhadap seluruh data *formant* yang ada sehingga pada akhirnya akan

ditemukan jumlah anggota kelompok yang sama dari *centroid* yang berbeda jumlah *formant*-nya . Pada akhirnya diperoleh suatu pengelompokan yang membedakan antara kelompok Dewasa dengan kelompok Anak – anak . Berikut adalah tabel data *formant* yang diperoleh dari hasil *training* :

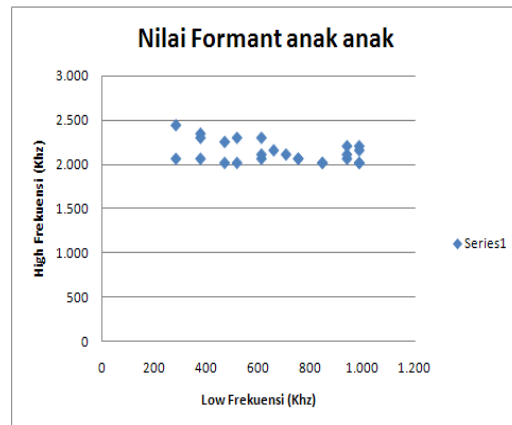
a) Gambar hasil nilai Formant Dewasa



Gambar 19 nilai Formant Dewasa

Dari Gambar 19 terlihat bahwa nilai Formant2 object dewasa yang diambil data rata rata nilainya diatas 2500

b) Tabel Nilai Formant anak – anak :

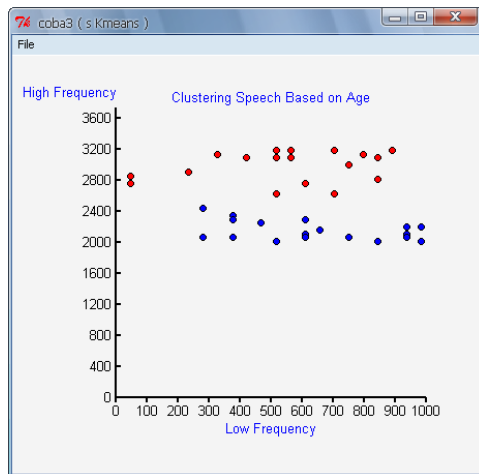


Gambar 20 Tampilan titik nilai Formant anak anak

Dari Gambar 20 dapat disimpulkan pada nilai Formant object anak anak, rata rata nilai Formant 2 dibawah 2500

Untuk Lebih Jelasnya akan dijelaskan dengan gambar 21 dan rincian sebagai berikut :

a) **Gambar Hasil Pengklusteran Dewasa dan Anak – anak**



Gambar 21 Hasil Pengklusteran Dewasa dan Anak – anak dengan menggunakan tcl/snack.

Keterangan Gambar

- Kluster atas atau nilai centroid 1 (Dewasa)
- Kluster bawah atau nilai centroid 2 (Anak-anak)

Dari pengklusteran didapatkan 2 Nilai Centroid, yakni :

- a. Centroid 1 : 564.3 2987.8
- b. Centroid 2 : 667.6 2139.35

b) **Pengujian pengenalan system :**

Pada pengujian Sistem dapat dilihat pada tabel 4 dengan penjelasan :

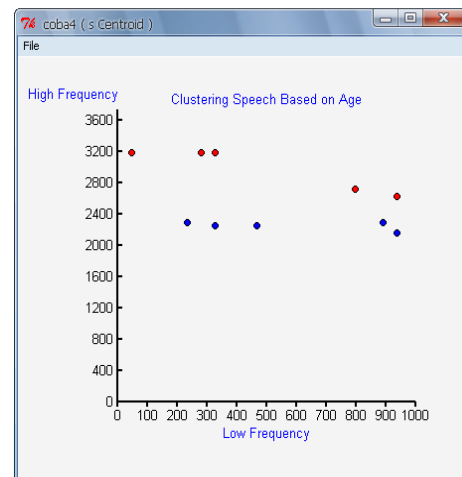
1. Data yang diambil dalam pengujian pengenalan system sebanyak 10 data yang terdiri dari 5 Dewasa dan 5 anak-anak.
2. Dimana batas threshold dari Dewasa adalah diatas 2500 sedangkan pada anak-anak dibawah 2500. Diambil dari perwakilan nilai formant masing-masing kelas clustering.
3. Centroid telah ditentukan yakni dari hasil pengambilan data secara random sebelumnya, yakni pada saat pengambilan dan pengamatan dari 50 data sample (25 Dewasa dan 25 Anak-anak)
4. Sehingga dari sini pada saat pengujian pengenalan system akan dinyatakan benar apabila data yang dimasukkan masih dalam batas threshold dan centroid masing-masing cluster.
5. Adapun nilai centroid yang telah ditentukan adalah sbb :

- a. Centroid 1 : 564.3 2987.8
- b. Centroid 2 : 667.6 2139.35

Tabel 4 Pengujian pengenalan Sistem

Nama Object perekam	Usia (tahun)	Nilai Formant 1	Nilai Formant 2	Keterangan
Anhar	25	328.000000	3187.000000	Benar (Dewasa)
Dedi	20	281.000000	3187.000000	Benar (Dewasa)
Dito	22	46.000000	3187.000000	Benar (Dewasa)
Kamal	23	937.000000	2625.000000	Benar (Dewasa)
Andin	19	796.000000	2718.000000	Benar (Dewasa)
Rido	12	234.000000	2296.000000	Benar (Anak-anak)
Robi	10	328.000000	2250.000000	Benar (Anak-anak)
Lita	11	468.000000	2250.000000	Benar (Anak-anak)
Ajeng	13	937.000000	2156.000000	Benar (Anak-anak)
Putri	14	890.000000	2296.000000	Benar (Anak-anak)

6. Dari pengujian diatas dapat disimpulkan system dapat mengenali 100 % object sebagai cluster Dewasa atau anak-anak.
7. Untuk lebih jelasnya dapat dilihat pada gambar 22 dimana warna merah adalah cluster dewasa dan warna biru adalah cluster anak-anak.



Gambar 22 hasil pengujian pengenalan system

5. Rencana Selanjutnya

Hasil pada paper ini masih merupakan hasil sementara, sehingga perlu dilakukan pembahasan lebih lanjut mengenai analisa tentang Clustering suara berdasarkan Usia dengan menggunakan metode K – Means. Atau bahkan juga bisa menggunakan metode clustering yang lain pada penelitian selanjutnya.

5.1 Kesimpulan

Berdasarkan pada hasil pengujian dan analisa terhadap hasil yang didapatkan, maka dapat diambil suatu kesimpulan yaitu :

1. Dari hasil pengujian dan analisa kita dapat membedakan suara anak dan dewasa dari nilai formant 2 nya dimana dari hal ini nilai formantnya 2 nya untuk dewasa lebih dari 2500 sedangkan pada anak anak dibawah 2500
2. Sedangkan untuk nilai formant 1 nya baik pada dewasa maupun anak anak menyebar.
3. Clustering dengan menggunakan K means menghasilkan 2 nilai centroid dimana nilai centroid tersebut akan menentukan kelas clustering.
4. Tingkat keberhasilan sistem 100% dengan jumlah data yang diujikan 10 (5 Dewasa dan 5 Anak anak) untuk pengujian secara *offline*.

5.2 Saran

Mengingat banyaknya kekurangan dalam proyek akhir ini, maka sebagai saran yang dapat diberikan untuk memperbaiki kinerja sistem yang dapat dilakukan pada penelitian selanjutnya :

1. Otomatis sistem, dimana pengguna tidak perlu menggunakan piranti tambahan untuk menjalankan *software* pengenalan suara.
2. Kecepatan sistem, Waktu pemrosesan sinyal suara untuk mendapatkan ciri atau parameter dipersingkat sehingga sistem pengenalan suara dapat bekerja secara *real time*.
3. Pengambilan sample yang lebih banyak lagi agar bisa didapatkan hasil pengenalan suara yang lebih akurat.
4. Tingkat pengurangan noise pada waktu perekaman suara yang semakin baik.
5. Untuk penelitian berikutnya berikutnya disarankan dicoba untuk jumlah data yang lebih besar.
6. Disarankan agar dilakukan penelitian lebih lanjut dengan menggunakan Metode K

means yang lain maupun dengan menggunakan metode cluster yang lain.

6. Daftar Pustaka

- [1] Achmad Solichin, “Pemrograman Bahasa C dengan Turbo C”, Copyright © 2003 IlmuKomputer.Com.
- [2] John. G. Proakis, Dimitris. G. Monolakis, “Digital Signal Processing: principles, algorithms, and application”, Prentice Hall, Inc, New Jersey, 1995.
- [3] Budi S Tri, “Praktikum Aplikasi DSP”, PENS – ITS, Surabaya, 2008.
- [4] Huda Miftahul, Bima, “Pelatihan Tcl/Tk”, PENS – ITS, Surabaya, 2005.
- [5] Eka.K Yesika,” Pembuatan Software Pembuka Program Aplikasi Komputer Berbasis Pengenalan Suara ”, PENS – ITS, Surabaya, 2006.
- [6] Solihah Maratus, “Aplikasi Pengenalan Suara untuk Sistem Akses Buku Perpustakaan”, PENS – ITS, Surabaya, 2007.
- [7] Yudi Agusta,” K-Means – Penerapan, Permasalahan dan Metode Terkait” Jurnal Sistem dan Informatika Vol. 3, 47-60, STMIK STIKOM BALI Denpasar Bali, (Pebruari 2007).